

Hello, my name is Walter Sujansky and I am the Chief Technology Officer of the California Joint Replacement Registry. I would like to thank the Health I.T. Policy Committee and the Department of Health and Human Services for the opportunity to testify today on the very important topic of clinical data quality and its implications for the secondary uses of EHR data .

The California Joint Replacement Registry, or CJRR, is a recently developed registry of patients who have undergone total hip or total knee replacement in California. The goal of the CJRR is to aggregate data from provider organizations and patients across the state into a single consolidated and standardized database. The purpose of the registry is to support (1) scientific analysis of best practices in joint replacement therapy, (2) self-assessment and benchmarking by providers with respect to quality metrics, (3) more convenient quality reporting to private, state and national entities, and (4) post-market surveillance of total-joint prostheses. The registry, which is currently funded by the California HealthCare Foundation and managed by the Pacific Business Group on Health, is in its second year of operation. To date, it has seven high-volume hospitals participating and has collected data on over 2500 joint-replacement cases.

The CJRR is a “level 3” registry, meaning that it contain not only information about the joint-replacement surgeries themselves and the implanted devices, but also data about the patients’ risk factors, surgical complications, clinical outcomes, and patient-reported functional status. The efficient and scalable collection of this information requires the harvesting of data from existing information systems at the participating organizations. These systems include

- Billing systems
- Electronic Health Record Systems
- Surgical Information Systems
- Laboratory Information Systems
- Clinical Data Warehouses
- Local Joint-Replacement Registries

Like most registries that re-purpose data originally collected for other uses, the CJRR has encountered a number of data-quality challenges. I’ll mention three during my prepared comments.

1. Most hospitals contributing data to the CJRR rely heavily on billing data to identify total joint-replacement cases and to report relevant patient risk factors and surgical complications to the registry. Traditionally, billing data have proven less than 100% accurate in reflecting the clinical condition of patients. However, for the risk factors and complications data collected by the CJRR, a sensitivity and specificity of 80-100% has been achieved to date, per preliminary chart audits . We speculate that this relatively high accuracy is achieved due to the rigorous coding processes in place at inpatient facilities, given the financial implications of accurately capturing patient co-morbidities and treatment complications. However, we are continuing to measure and, where necessary and possible, improve the accuracy of these billing data.
2. Hospitals contributing data to the CJRR also access their clinical information systems for data such as patient height and weight, surgical approach, laboratory results, and certain preventative interventions . In our experience, these data sources are more difficult to draw upon successfully. In certain cases, this is because the data are stored as free text, rather than the structured and coded representations that lend themselves to automated retrieval and reporting. In other cases, the data are represented using local or proprietary coding systems that do not readily correspond to the standard coding required by the registry. In particular, these issues complicate the mapping of local data to standard definitions of quality measures,

and many of the participating sites are unable to report these measures to the CJRR in an automated way.

3. A third, more general data-quality challenge is the limitation in access that certain hospitals have to the data in their own information systems. This issue arises when the vendors of proprietary systems limit the hospitals' direct access to the data stored in these systems. In certain cases, the limitations are technical, in that no useful query facilities and/or documentation exists. In other cases, the limitations are due to explicit policies of the vendors, which are enforced through contractual mechanisms. The result is that hospitals must involve the vendors in many data extraction and export processes, and are unable to leverage their internal personnel or third-party consultants. When the hospitals' or the vendors' resources are limited, this constraint can slow or impede the collection of data for secondary uses, such as the CJRR. We have experienced this first hand.

The CJRR is addressing these various data-quality and data-collection challenges through various approaches. These approaches include data-quality audits via chart-reviews at all of the participating sites, automated surveillance of the CJRR database for patterns suggesting missing or inaccurate data, and enlistment of information-system vendors to develop data-collection mechanisms for the CJRR that can be re-used across their customer bases. I look forward to further discussing our experiences and those of others during today's conversation.